The Increasing Influence of Risk Assessment on Forensic Patient Tribunal Decisions

N. Zoe Hilton
Waypoint Research Institute

Alexander I. Simpson
Centre for Addiction and Mental Health

Elke Ham
Waypoint Research Institute

Version Post-print/accepted manuscript


Publisher's Statement ©American Psychological Association, [2016]. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article is available, upon publication, at: DOI 10.1037/ser0000068.

How to cite TSpace items

Always cite the published version, so the author(s) will receive recognition through services that track citation counts, e.g. Scopus. If you need to cite the page number of the author manuscript from TSpace because you cannot access the published version, then cite the TSpace version in addition to the published version using the permanent URI (handle) found on the record page.

This article was made openly accessible by U of T Faculty. Please tell us how this access benefits you. Your story matters.
The Increasing Influence of Risk Assessment on Forensic Patient Tribunal Decisions

N. Zoe Hilton
Waypoint Research Institute

Alexander I. Simpson
Centre for Addiction and Mental Health

Elke Ham
Waypoint Research Institute


Author Note

N. Zoe Hilton, Department of Psychiatry, University of Toronto and Waypoint Research Institute, Penetanguishene; Alexander I. Simpson, Clinician Scientist and Chief of Forensic Psychiatry, Centre for Addiction and Mental Health (CAMH), and Department of Psychiatry, University of Toronto; Elke Ham, Waypoint Research Institute, Penetanguishene.

We would like to thank Chelsea Turan, Jenna Rutherford, Oleg Belanovsky, and Sonja Dey for research assistance, and Daryl Kroner and Jeremy Mills for helpful comments on the manuscript.

Correspondence concerning this article should be sent to N. Zoe Hilton, Ph.D., Department of Psychiatry, University of Toronto, and Waypoint Research Institute, Penetanguishene. 500 Church Street, Penetanguishene, ON, L9M 1G3 Canada. Email: zhilton@waypointcentre.ca; zoe.hilton@utoronto.ca
Abstract

Previous studies of mental health tribunal decisions about forensic patients indicate some changes over time in the use of empirically supported risk factors. Our aim was to investigate whether, in more recent cases, risk assessment instruments were cited by the tribunals or the clinicians making recommendations to them, and whether there was evidence of an association between risk assessment results and either clinical recommendations or tribunal dispositions. Among review board hearings held in 2009-2012 pertaining to 63 different maximum security patients found not criminally responsible on account of mental disorder in Ontario, Canada, dispositions were most strongly associated with psychiatrists’ testimony, consistent with previous studies. Dispositions were, however, associated with the scores on the Violence Risk Appraisal Guide (VRAG), such that transferred patients had a lower risk of violent recidivism than detained patients. An association between clinical opinions and risk assessment results was also found to be significant, and significantly larger than in previous research. There was no evidence that risk assessment was cited selectively in higher risk cases or when scores were concordant with the tribunal decision. This research may provide a baseline for studies of the effect of 2014 legislations introducing a high-risk designation for forensic patients in Canada. We recommend further efforts to measure the effect of non-pharmacological treatment participation and in-hospital security decisions on tribunal decision-making.

Key Words: forensic decision-making, forensic patients, review board, tribunal, risk assessment
The Increasing Influence of Risk Assessment on Forensic Patient Tribunal Decisions

The development and validation of instruments to assess the risk of violent offending has greatly expanded in the past 25 years (e.g., Harris, Rice, Quinsey, & Cormier, 2015; Mills, Kroner, & Morgan, 2011). The emerging literature on how risk assessment influences forensic decision making, however, is sparse in comparison (e.g., Hilton, Scurich, & Helmus, 2015). Mental health tribunals appear to be largely dependent on the clinicians’ recommendations for the most appropriate security disposition for forensic psychiatric patients (e.g., Hilton & Simmons, 2001; McKee et al., 2007; Wilson, Crocker, Nicholls, Charette, & Seto, 2015) similar to comparable judicial and conditional release decisions (e.g., Blais, 2015; Guy, Kusaj, Packer, & Douglas, 2015; Konečni & Ebbesen, 1984). Nevertheless, recent research shows that mental health tribunals refer to empirically supported risk factors such as previous violence, substance use, and personality disorder in the reasons given for their dispositions (Wilson et al., 2015).

The purpose of the present research was to investigate whether recent reasons for dispositions refer to risk assessment instruments themselves, in a sample of forensic patients held under conditions of maximum security care. We aimed to detect any trend in the influence of risk assessment tools in mental health tribunal decisions compared with previous research.

In Canada, decisions regarding the disposition of forensic patients to treatment in secure hospital or other locations are conducted under the auspices of a body known as the Review Board, legislated by the Criminal Code of Canada. Prior to 1992, tribunals served an advisory function, reporting to the Lieutenant General. Under section 672 of the Criminal Code, each province has established a Review Board, of at least five members including at least one psychiatrist and at least one judge, to make or review dispositions for those found not criminally responsible on account of mental disorder or unfit to stand trial. Dispositions may include conditional or absolute discharge to the community, or detention in a hospital which in practice
may include a general psychiatric unit or a medium or maximum security forensic unit. Case law (Winko v. British Columbia, 1999) resulted in the onus of proof being changed from the patient demonstrating a lack of risk to public safety in order to obtain discharge, to the Review Board demonstrating an existing risk in order to support not discharging the patient. In making its disposition, the Review Board is required to take into account “the safety of the public, which is the paramount consideration, the mental condition of the accused, the reintegration of the accused into society and the other needs of the accused” (Criminal Code s. 672.54). The paramount consideration of risk to others, combined with the need to also consider aspects of mental health, has led to a small but growing area of research into the relative influence of measurable risk compared with other variables such as psychiatric presentation and treatment.

In the first study of dispositions since Canada’s current Review Board model was established, Hilton and Simmons (2001) reported that dispositions were strongly related to recommendations from the clinical team and from the senior clinician testifying at the tribunal, but not to scores on an actuarial risk assessment tool, the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993; Harris et al., 2015). Clinician testimony itself was unrelated to the VRAG. Subsequently, McKee, Harris, & Rice (2007) studied dispositions concerning 104 patients in the same population resulting from hearings that took place in 2000-2003. Decisions were less strongly related to clinician testimony and team recommendation, and clinician testimony was now significantly positively related to the VRAG. In a national study of 6743 Review Board decisions across Canada up to the end of 2008, including 2185 taking place in Ontario, decisions to detain versus discharge patients were significantly associated with scores on the Historical Clinical Risk Management-20 (HCR-20; Douglas, 2014; Webster, Douglas, Eaves, Hart, 1997), a structured risk assessment tool (Crocker, Nicholls, Charette, & Seto, 2014),
and items from both the HCR-20 and the VRAG were frequently mentioned by clinicians and Review Boards (Wilson, Crocker, Nicholls, Charette, & Seto, 2015). Similar findings have been observed for clinical and correctional reports to courts (e.g., Storey, Watt, & Hart, 2015).

The findings across these studies suggest that forensic decision making has become more associated with measured risk of recidivism over time. Despite this apparent trend, tribunals and courts continue to give more consideration to variables such as major mental illness, treatment, and remorse than to risk assessment scores (e.g., Turner, Boccaccini, Murrie, & Harris, 2015; Wilson et al., 2015). For example, the HCR-20 as a complete instrument was cited only in 9% of Review Board dispositions, and the VRAG, in 8% (Wilson et al., 2015). A statistically reliable association between dispositions and actuarial risk of recidivism, moreover, has yet to be observed. Furthermore, Review Board decisions were associated with whether a structured risk assessment tool was used (Croker et al., 2014), raising the question of whether risk assessment is used disproportionately to support decisions to detain patients in hospital.

We conducted the present study in order to examine whether the positive trend of an association between measured risk of recidivism and tribunal decision making continued in a more recent sample of hearings conducted in 2009-2012. Because much previous research has focused on the VRAG, and because it was originally developed for a forensic assessment population such as the one in the present study and it provides a precise numeric estimate of risk, we used the VRAG as our main measure of risk. We examined whether there would continue to be a positive association between the VRAG score and clinical opinions and whether the effect would be larger than in the past. We also explored the effect of the VRAG and other risk assessment tools being cited in the reasons given for dispositions, and in particular whether tools were more likely to be cited when they were concordant with the decision. We expected that the
HCR-20 would also be used and anticipated a similar positive association between it and clinical opinions.

We aimed to extend previous research by examining the effect of patients’ participation in therapy. Treatment program completion has been related to security level changes within a forensic hospital, such that those who made less progress in treatment had more increases in security (Davoren et al., 2012). In research examining the factors associated with clinical opinions regarding transfer from secure hospital settings, though, patients’ compliance with medication has been considered but not participation in other forms of treatment (e.g., Croker et al., 2015). Hilton and Simmons (2001) were unable to measure participation in therapeutic activities reliably, and McKee et al. (2007) measured only whether patients refused treatment or showed insight into their illness. The absence of non-medical therapy in research on tribunal decision making limits evaluation of the relative influence of public safety needs and treatment and rehabilitation needs, and thus of the fulfillment of these considerations as required by the Criminal Code. In the present study, therefore, we measured whether patients attended and completed a variety of therapeutic activities.

**The Present Study**

We examined Review Board decisions made in 2009-2012 concerning patients found not criminally responsible on account of mental disorder (NCR) and detained in Ontario’s maximum security hospital. We examined whether the Violence Risk Appraisal Guide or other assessments were cited in the reasons for disposition, and whether dispositions were related to the assessment scores, as well as to the psychiatrist testimony and clinical team recommendation. We hypothesized:
1. Dispositions to transfer patients would show a strong positive association with clinical opinions (i.e., psychiatrist testimony and clinical team recommendation), consistent with previous research;

2. Dispositions and clinician opinions would show a significant association with actuarial risk for violent recidivism as measured by the VRAG score, in that patients with higher risk scores would be more likely to be detained, and those with lower risk scores would be more likely to be transferred, extending the trend in recent research;

3. Dispositions and clinician opinions would show a significant association with non-actuarial structured risk assessment as measured by HCR-20 scores and summary risk judgments, such that patients with higher risk scores and judgments would be more likely to be detained, and those with lower HCR-20 risk scores and judgments would be more likely to be transferred.

We also examined whether dispositions and clinical opinions were associated with participation in therapeutic activities, and explored whether the VRAG was more likely to be cited when it was concordant with the Review Board disposition.

**Method**

**Sample**

The present sample was drawn from a longitudinal study of men admitted to Ontario’s maximum security forensic hospital division, a 160-bed unit serving the entire province housing male patients only. At the time of the present study, data had been collected on admission and after one year for all men admitted in 2009 through 2011. We included only men who were admitted after a finding of NCR or who were found NCR during the first year. The unit of study was the Review Board hearing held for each NCR patient at this hospital during the first year. In
the event that two hearings were held for a patient within his first year \( n=5 \), which was usually because of unresolved issues at the first hearing or to allow time for stability of patients’ clinical presentation, we recorded the second hearing. The resulting 63 hearings pertained to 63 different patients and took place between May 2009 and July 2012.

**Procedure**

Data were coded from patients’ medical records at two time points. Shortly after the admission assessment was completed, research coders read the psychosocial history report, criminal history, and other information pertaining to the index offense (for which the patient was eventually found NCR) and pre-index history in order to score the VRAG and diagnosis. They also noted the PCL-R score if it was documented during the assessment. After a year since the admission, one of four coders reviewed clinical progress notes, the hospital report to the Review Board, and the Review Board reasons for disposition in order to code the remaining variables, including the HCR-20 if it was documented.

**Measures**

*Review Board Dispositions.* The principal dependent measure was the outcome of the Review Board hearing. Review board dispositions were measured in one of five categories intended to represent an ordinal scale: 0 (absolute discharge), 1 (medium or minimum security with community access), 2 (medium security with no community access), 3 (medium security for temporary period other than for assessment), 4 (remain in maximum security); however, in practice only one patient received an absolute discharge and all who were ordered to a medium secure hospital were given community access; therefore, we treated this variable as a binary outcome of detain versus transfer or discharge.
Clinical Opinions. A senior clinician, normally the patient’s attending psychiatrist, presented testimony at the hearing and his or her viva voce testimony was recorded in the subsequent reasons for disposition. We coded the psychiatrist’s testimony from this document as 1 (recommends transfer), 2 (no explicit recommendation but testimony favors transfer; e.g., testimony focuses on treatment success), 3 (undecided), 4 (no explicit recommendation but testimony favors detention; e.g., reservations expressed about ability to succeed in a lower security environment), 5 (recommends detention in maximum security). In our analyses we first treated psychiatrist testimony as an independent measure and examined its association with dispositions, then as a dependent measure as we examined factors associated with the testimony. In two cases, there was no evidence in the reasons for disposition of a recommendation made by the psychiatrist.

The multi-disciplinary clinical team normally held a meeting prior to the hearing to form its opinion and make a recommendation. This recommendation was recorded in the hospital report to the Review Board. We coded the team’s recommendation as 1 (unanimous to transfer), 2 (majority opinion to transfer), 3 (split opinion), 4 (majority to detain), 5 (unanimous to detain). In nine cases, the team recommendation was not reported.

Assessment Instruments. Our main measure of risk was the Violence Risk Appraisal Guide (VRAG; Harris et al., 1993; Harris et al., 2015). The VRAG would normally be scored by a clinical psychologist prior to a tribunal hearing and a psychological report placed on the medical record. We expected to code the VRAG based on a clinical assessment documented in the medical record, but it was not documented in approximately one third of cases; therefore, we scored the VRAG for every case following the procedure in the scoring manual, prorating in cases with up to four missing items. In one case, there was insufficient information to code the
VRAG. One VRAG item is the PCL-R total score or substitution with the Child and Adolescent Taxon Scale (CATS; Harris et al., 2015). We used the PCL-R as scored by clinicians and documented in the clinical record where available (68%) or substituted the CATS score after prorating it out of 40.

There was no separate report on file for the HCR-20 assessment. We recorded scores for the HCR-20 (range from 0 to 40) and any summary risk judgments when they were cited by the Review Board in its reasons for disposition or by the clinicians in the hospital's report to the Review Board. Historical, Clinical, and Risk sub-scales of the HCR-20 were rarely cited separately so we did not examine their unique contributions to clinical opinions or dispositions.

We also coded whether the VRAG, PCL-R, and HCR-20 were cited by the Review Board or clinicians.

**Participation in Therapeutic Activities.** We coded participation in non-pharmaceutical therapy within the first year of admission. We coded participation separately for four types of therapeutic activities defined by their structure and content. Organized, scheduled activities that require or build skills (“skills therapy”) mostly consisted of education or doing piecework or skilled work. Structured, scheduled group sessions following a manual or prepared curriculum (“group therapy”) and structured, scheduled individual sessions following a manual or prepared curriculum (“individual therapy”) typically concerned symptom management or a potentially criminogenic need such as anger management or substance use. A fourth activity category included unstructured, individual sessions offering supportive counseling, spiritual counseling, or other therapy (“counseling”). For each type of therapeutic activity, we measured participation as 0 (declined or was not offered), 1 (occasional participation; irregular attendance or started but dropped out of a structured program), or 2 (regular participation; consistent attendance or
completed a structured program). In addition, we coded refusal to participate in nonmedical therapy on a five-point scales labeled at 1 (no problem), 3 (moderate problem; persistent refusal), and 5 (severe problem; pervasive, disruptive, acrimonious pattern of refusal; e.g., becomes aggressive when therapy is offered). Similarly, we coded lack of insight into mental illness on a five-point scale labelled at 1 (no problem), 3 (moderate problem; reported to have no insight into mental illness), and 5 (severe problem; pervasive, disruptive, acrimonious pattern of denying problems).

**Inter-Rater Reliability**

Inter-rater reliability was assessed by comparing researchers’ independent coding of cases drawn from the longitudinal study. The VRAG score was based on ten cases and yielded a single rater, absolute agreement intraclass correlation (ICC) of .99. Agreement on remaining variables was based on a total of 43 cases and ICCs exceeded .96 for the scaled measures of disposition, psychiatrist testimony, and team recommendation. For categorical measures of participation in treatment variables agreement ranged from 73% to 89%. Agreement was lower for refusal to participate in therapy, ICC = .33, and insight into mental illness, ICC = .53.

**Analytic Procedure**

We compared detained and transferred patients on key variables using chi-squared tests for binary variables and t-tests for remaining variables. We also examined the correlation between the VRAG score and dispositions, clinical opinions, and assessments.

We tested hypothesis 1, that tribunal dispositions would show a strong positive association with psychiatrist testimony and clinical team recommendations, and hypotheses 2 and 3, that dispositions would show a positive association with VRAG and HCR-20 scores, using point-biserial correlations in order to make comparisons with previous research. We also used
the ROC area under the curve statistic (AUC) as a measure of effect size that is less prone to attenuation due to base rates (Babchishin & Helmus, in press); AUCs greater than .71 were considered large effect sizes (Rice & Harris, 2005). We anticipated testing hypothesis 3 using chi-squared tests of the association between dispositions and HCR-20 summary risk judgments. We further tested hypotheses 1 and 2 using binary logistic analysis of tribunal decision entering psychiatric testimony, clinical team recommendation, and VRAG score. We tested whether clinician opinion would be more strongly associated with VRAG score than reported in previous research, using the Fisher $r$-to-$Z$ score transformation (calculated online at http://vassarstats.net/rdiff.html). We used chi-squared tests to examine the association between opinions and each therapy category (skills-building, group therapy, individual therapy, counselling) and AUCs to examine the association between opinions and treatment refusal and insight into illness.

We explored whether the VRAG was more likely cited when the score was concordant with dispositions. We defined concordance as either decisions to detain when the VRAG score was 0 or higher (i.e., numeric risk categories 5 to 10) or decisions to transfer when the VRAG score was -1 or lower (i.e., numeric risk categories 1-4). The cut off of zero was chosen because the mean VRAG score in its construction was close to 0 and because this score demarcates the lower bound of the middle of nine risk categories (Harris et al., 1993; Harris et al., 2015). We defined discordance as decisions to detain when the VRAG score was -1 or lower (numeric risk categories 1-4) or dispositions to transfer when the VRAG score was 0 or higher (numeric risk categories 5 to 10). We then used a chi-squared test of the association between disposition and concordance.
We explored whether clinical opinion was associated with therapy participation, refusal to participate in therapy, and patients’ insight into their mental illness using bivariate statistics.

**Results**

There were 63 patients with Review Board hearings pertaining to their NCR status within the first year of admission. The mean age of these men was 29.9 years ($SD = 10.8$). Most (70%) were charged with a violent offense, ranging from assault to murder; others were charged with nonviolent offenses including arson, theft, harassment, break and enter, and mischief. The most common primary diagnosis was schizophrenia (54%), followed by other psychosis (29%), mood disorder (8%), personality disorder (5%), developmental disorder (3%), and substance related disorder (2%). The average time between the first admission to maximum security for the index offense and the hearing was 6.58 months ($SD = 3.32$) and most patients (65%) spent at least one year in the hospital.

The VRAG was referred to in the hospital report to the Review Board or in the Review Board reasons for disposition in 45 (71%) cases, in which the numeric risk category was cited. The HCR-20 was referred to in 11 (18%) cases, in which the numeric total score was cited in 8 cases, the summary risk judgment was cited in two cases (both “high risk”), and no score or judgment was given in one case. When the VRAG was cited, the HCR-20 was also cited in 21% of cases. Correlations among disposition, opinions, and assessments are shown in Table 1.

**Review Board Dispositions and Clinical Opinions**

We hypothesized that dispositions to transfer patients would show a strong positive association with psychiatrist testimony and clinical team recommendations. Consistent with hypothesis 1, dispositions were positively associated with both psychiatrist testimony and clinical team recommendation (Table 1) with a large effect size, psychiatrist AUC = .946 (SE =
.034), 95% CI = [.880, 1.00], clinical team AUC = .929 (SE = .040), 95% CI = [.851, 1.00]. That is, where clinicians supported transfer to lower security, the disposition was most often to transfer. Disagreements between the psychiatrist testimony and tribunal disposition were rare; two involved a decision to detain patients recommended for transfer, and one involved the reverse.

**Dispositions, Opinions, and VRAG**

We hypothesized that dispositions and clinician opinions would show a significant association with VRAG scores. Consistent with hypothesis 2, disposition decisions were associated with the VRAG score (Table 1) with a large effect size, AUC = .714 (SE = .068), 95% CI = [.581, 847]. That is, patients who were detained had a higher risk of violent recidivism according to the VRAG than detained patients (Table 2). Table 2 also shows that a similar pattern was found for PCL-R score.

There was a moderate association between the psychiatrist recommendation to transfer and the VRAG score, AUC = .708 (SE = .069), 95% CI = [.573, 843]. The correlation, \( r (df = 59) = .33 \), was significantly larger than the comparable statistic reported by Hilton and Simmons (2001), -.02, \( Z = 2.39, p = .017 \), and larger but not significantly so than that by McKee et al. (2007), .23, \( Z = 0.66, p = .511 \). This result also lends support to hypothesis 2.

We further tested hypotheses 1 and 2 using binary logistic analysis of disposition to detain, entering psychiatric testimony, clinical team recommendation, and VRAG score. Only psychiatrist testimony yielded a significant contribution, \( B = 1.688 (SE = .762) \), Wald = 4.91, \( p = .027 \), \( \text{Exp} (B) = 5.41 \), supporting hypothesis 1 that the disposition would be associated with clinical opinion (psychiatric testimony). The model resulted in 96% correct classification: 100%
transfer decisions, 93% detain decisions. Neither team recommendation nor VRAG made a unique contribution.

**Dispositions, Opinions, and HCR-20**

We hypothesized that dispositions and clinician opinions would show a significant association with HCR-20 scores and summary risk judgments where cited. Consistent with hypothesis 3, dispositions were associated with HCR-20 scores (Table 1). That is, in the eight cases where the HCR-20 was reported (6 patients transferred, 2 detained), detained patients had higher HCR-20 scores than detained patients, $M = 31.33$ ($SD = 5.39$) vs. $M = 15.50$ ($SD = 10.61$), 95% CI of the difference = [2.73, 28.93]. This result is presented with caution because it is based on so few cases, and associations with risk summary judgments were not tested because of their low occurrence.

**Clinical Opinions and Therapy Participation**

The most participation occurred for skills therapy (65% at least occasional participation) followed by group therapy (56%) and counseling (30%). Only two patients (3%) participated in individual therapy, one of whom was recommended for transfer. Patients who participated in group therapy were less likely than other patients to be recommended for transfer by psychiatrists (34% vs. 70%), and the level of participation (from none to regular) was associated with clinical opinions in an unexpected direction (Table 1) including psychiatrist recommendation to detain, AUC = .686 ($SE = .069$), 95% CI = [.550, .822], and clinical team recommendation to detain, AUC = .706 ($SE = .072$), 95% CI = [.565, .842]. Correlations between group therapy and clinical opinions were un-reduced when controlling for length of stay, meaning this counterintuitive finding cannot be explained by therapeutic activities being more available to patients who were hospitalized longer. Effect sizes for skills therapy and counseling
were not statistically significant, AUCs .375 to .564. Therapy participation was generally unrelated to medication adherence, AUCs .332 to .499, as were treatment refusal and lack of insight (Table 1).

**Review Board Citation of Risk Instruments.** There was no evidence that the VRAG was more likely to be cited in decisions to detain rather than to transfer, nor that the VRAG was more often cited if the score was relatively high (Table 1). Most decisions were concordant with the VRAG score. There was no evidence of the VRAG being cited preferentially when the score was concordant with the decision (73%) than when it was not (72%), $\chi^2 (df = 1) = .007, p = .933$.

**Discussion**

In the present study of Ontario Review Board tribunal decisions concerning 63 men found not criminally responsible on account of mental disorder, dispositions were most strongly related to clinical opinion, as in previous research. The psychiatrist’s testimony at the Review Board hearing with respect to recommended security was the sole statistical contributor to the disposition, and actuarial risk assessment made no independent contribution. This result is very consistent with previous research with a similar population (e.g., Hilton & Simmons, 2001; McKee et al., 2007). Compared with earlier studies, however, in which risk assessment had little or no influence on the clinical opinion, psychiatrists’ recommendations were now more strongly influenced by risk assessment. There was also preliminary evidence that dispositions were also related to scores on a structured professional judgment instrument, consistent with recent work in a larger sample (Crocker, Nicholls, Charette, & Seto, 2014). There was no evidence that risk assessment was cited selectively in higher risk cases or when it was concordant with the tribunal decision. Together, these results provide evidence of a trend such that decisions regarding the placement of forensic patients appear to be increasingly influenced by violence risk assessment.
Participation in non-pharmacological treatment was common, with over half of patients having at least some level of participation in skills building and structured group therapy. We obtained counter-intuitive findings whereby patients who participated in therapeutic activities were more likely to be recommended to remain in the secure hospital, and this relation was not explained by patients’ length of stay. In an effort to improve reliability, we defined our measures in terms of clear criteria for the structure and nature of the activity rather than the intended therapeutic goal. As a result, we were unable to investigate alternative explanations such as whether clinicians considered greater security to be required for patients in need of (and consequently accepted into) non-pharmacological treatment.

**Methodological Limitations and Strengths**

The present sample might not be directly comparable to those of Hilton and Simmons (2001) and McKee et al (2007) because we focused on hearings held in the first year after patients were found not criminally responsible, whereas some of the cases in previous research had already experienced lengthy hospitalization. We did not attain acceptable inter-rater reliability on refusal to participate in therapy or lack of insight into mental illness, which could explain the small effects obtained for these variables and limits the comparability of these findings with other studies.

This study had a small sample which likely limited the detectable effects of risk in multivariate analysis. Dispositions to detain or transfer were approximately equal in the present study, whereas in previous studies the base rate of transfers differed more substantially (35% and 38%) which could have contributed to their smaller correlation coefficients and exaggerated the apparent increase in the influence of risk assessment. We provided AUC statistics, which are
more robust across varying base rates than correlations (e.g., Babchishin & Helmus, in press), in order to facilitate comparisons of effect sizes across future studies.

This study did not use a novel method, but replicated the approach of previous research. This allowed for a comparison of findings with similar research, and the discovery of an increasing influence of risk assessment on forensic decisions and clinical opinions in the present study compared with previous research. This study is the first to report a statistical association between actuarial risk of violent recidivism and both tribunal dispositions and clinical recommendations for the disposition of patients found not criminally responsible. In the context of previous research in this field, this finding is consistent with an increasing influence of risk assessment on forensic patient tribunal decisions.

**Implications for Practice and Directions for Future Research**

Criminal Code of Canada changes effected in 2014 increased the emphasis that its tribunals must place on patients’ risk to the safety of the public, by introducing a “high risk” category whereby patients newly found not criminally responsible on account of mental disorder may be detained in hospital initially for three years before review. The present findings suggest that risk assessment has already been having an influence on the detention of forensic patients in secure psychiatric hospital, prior to the availability of the high risk designation, equivalent to a moderate to large effect for forensic opinions and decisions.

While threat of safety to the public is only one of the considerations that the Review Board must make in its disposition orders, alongside the patients’ treatment and rehabilitation needs, changes in the wording of the statutory definition of ‘significant threat’ and making issues of public safety ‘paramount’ emphasize the need for appropriately structured assessment of risk and communication of that risk to Review Boards. A simple way to increase the use of risk assessments would be to ensure that a structured assessment report is routinely available,
compared with the present study in which we found that a VRAG assessment report was often not documented on the medical record. When the VRAG was used, the HCR-20 was also cited in one fifth of cases, and the scores were not significantly correlated. The use of multiple assessments raises challenges for interpreting and communicating potentially conflicting results (e.g., Seto, 2005), which could limit the influence of risk assessment on tribunal decisions.

Recommendations have been made for increasing the systematic integration of risk assessment in forensic decisions, such as embedding the assessment instrument into a policy whereby more restrictive dispositions would be assigned in accordance with assessed risk (e.g., Harris & Rice, 2015). This strategy can be combined with the communication of clinical and criminogenic needs and recommendations for evidence-based interventions (e.g., Cook, Murray, Amat, & Hart, 2014; Harris & Rice, 2015) which would support the Review Board’s other mandated considerations.

There is a need for research examining the use of the new Criminal Code high risk category and its relation to dispositions, such as whether patients deemed high risk are directed to be detained in maximum security. The association of the high risk designation with actuarial and structured professional judgment assessments of recidivism risk should also be examined, as well as any change in the adherence of dispositions to clinician recommendations. The hearings in the present study occurred during a time of relative stability in the incidence of new cases under the Ontario Review Board following a decade of changing legal trends and a concomitant surge in tribunal cases (Simpson, Penney, Seto, Croker, Nicholls, & Darby, 2014). Because the present study involved decisions made during this time of stability in the years just prior to the passing of the new legislation, future research could use the present study as a pre-legislation baseline.
There is an emerging literature on security levels within secure psychiatric care (e.g., Davoren et al., 2012; Miller-Isberner, Webster, & Gretenkord, 2007), in which risk factors, treatment participation, and other variables have been found to be related to increases or decreases of security or privilege within the hospital. Research combining this work with studies of factors related to transfers between institutions with different levels of security, and from hospital into the community, would permit the identification of patient trajectories within the forensic system. This could lead to an understanding of the variables and processes affecting the progress of patients from admission through to absolute discharge, and potential recognition of patient characteristics or circumstances likely to lead to long term institutional care. More exploration of methods for reliably measuring patients’ participation in non-pharmacological therapy, perhaps during a controlled time period during which patients are homogeneous for length of stay, will be important to a valid test of the role of therapy in patient progress.
References


Müller-Isberner, R., Webster, C. D., & Gretenkord, L. (2007). Measuring progress in hospital order treatment: Relationship between levels of security and C and R Scores of the HCR-


Table 1.

*Correlations among Study Variables*

<table>
<thead>
<tr>
<th>Variable</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
<th>11.</th>
<th>12.</th>
<th>13.</th>
<th>14.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Disposition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Psychiatrist Recommendation</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Team Recommendation</td>
<td>.85</td>
<td>.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. VRAG Cited</td>
<td>.13</td>
<td>.13</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. PCL-R Cited</td>
<td>.18</td>
<td>.18</td>
<td>.20</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. VRAG Score</td>
<td>.34</td>
<td>.33</td>
<td>.33</td>
<td>.21</td>
<td>.14</td>
<td>.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. PCL-R Score</td>
<td>.46</td>
<td>.43</td>
<td>.47</td>
<td>.09</td>
<td>.09</td>
<td>.37</td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. HCR-20 Score</td>
<td>.77</td>
<td>.71</td>
<td>.16</td>
<td>.82</td>
<td>.82</td>
<td>--</td>
<td>.29</td>
<td>.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Skills Therapy</td>
<td>.12</td>
<td>.11</td>
<td>.07</td>
<td>.24</td>
<td>.27</td>
<td>.13</td>
<td>.16</td>
<td>.22</td>
<td>-.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Group Therapy</td>
<td>.24</td>
<td>.35</td>
<td>.33</td>
<td>.15</td>
<td>.10</td>
<td>.24</td>
<td>.32</td>
<td>-.36</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Individual Therapy</td>
<td>.13</td>
<td>.13</td>
<td>-.15</td>
<td>.08</td>
<td>-.22</td>
<td>-.07</td>
<td>.17</td>
<td>.32</td>
<td>--</td>
<td>.01</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Counseling</td>
<td>-.31</td>
<td>-.28</td>
<td>-.19</td>
<td>.02</td>
<td>-.08</td>
<td>-.06</td>
<td>-.01</td>
<td>-.21</td>
<td>-.82</td>
<td>.12</td>
<td>.32</td>
<td>-.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Refusal to Participate</td>
<td>.02</td>
<td>.09</td>
<td>-.02</td>
<td>-.04</td>
<td>.05</td>
<td>-.22</td>
<td>-.23</td>
<td>-.12</td>
<td>.05</td>
<td>.04</td>
<td>.01</td>
<td>-.08</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>15. Lack of Insight</td>
<td>.09</td>
<td>.01</td>
<td>-.03</td>
<td>.01</td>
<td>-.04</td>
<td>-.17</td>
<td>-.02</td>
<td>-.12</td>
<td>.70</td>
<td>-.32</td>
<td>-.16</td>
<td>.27</td>
<td>-.22</td>
<td>.16</td>
</tr>
</tbody>
</table>

Note: Coefficients shown in italics did not achieve statistical significance at α = .05. HCR-20 was always cited where an HCR-20 score was known and therefore no correlation coefficient is shown. No offenders with an HCR-20 score participated in therapy and therefore no correlation coefficient is shown.
Table 2.

Clinical Opinions, Citation of Risk Assessment Instruments, and Risk Assessment Scores as a Function of Tribunal Decisions.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Review Board Disposition</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Detain 48% (n=30)</td>
<td>Transfer 52% (n=33)</td>
<td>95% CI for Group Difference</td>
<td>$\chi^2$ or $t$ statistic and $p$-value</td>
<td></td>
</tr>
<tr>
<td>Psychiatrist recommends transfer$^a$ (%)</td>
<td>3 (1)</td>
<td>94 (30)</td>
<td>0.79, 1.02</td>
<td>49.68, $p &lt; .001$</td>
<td></td>
</tr>
<tr>
<td>Team recommends transfer$^b$ (%)</td>
<td>4 (1)</td>
<td>90 (27)</td>
<td>0.71, 1.00</td>
<td>39.35, $p &lt; .001$</td>
<td></td>
</tr>
<tr>
<td>VRAG cited (%)</td>
<td>77 (23)</td>
<td>67 (22)</td>
<td>-0.13, 0.33</td>
<td>0.77, $p = .380$</td>
<td></td>
</tr>
<tr>
<td>PCL-R cited (%)</td>
<td>80 (24)</td>
<td>66 (21)</td>
<td>-0.08, 0.37</td>
<td>1.61, $p = .205$</td>
<td></td>
</tr>
<tr>
<td>HCR-20 cited (%)</td>
<td>30 (9)</td>
<td>7 (2)</td>
<td>0.04, 0.43</td>
<td>5.71, $p = .017$</td>
<td></td>
</tr>
<tr>
<td>VRAG score$^c$</td>
<td>12.63 (12.27)</td>
<td>4.21 (11.37)</td>
<td>2.41, 14.42</td>
<td>2.80, $p = .007$</td>
<td></td>
</tr>
<tr>
<td>PCL-R score$^d$</td>
<td>21.93 (8.22)</td>
<td>14.97 (5.39)</td>
<td>2.72, 1.18</td>
<td>3.32, $p = .002$</td>
<td></td>
</tr>
<tr>
<td>HCR-20 total score$^e$</td>
<td>31.33 (5.39)</td>
<td>15.50 (10.61)</td>
<td>2.73, 28.93</td>
<td>2.96, $p = .025$</td>
<td></td>
</tr>
</tbody>
</table>

Note: $^a$In two cases, no psychiatrist recommendation was reported ($n = 61$). $^b$In nine cases, no team recommendation was recorded ($n = 54$). $^c$VRAG = Violence Risk Appraisal Guide, scored by researchers from criminal and psychosocial history ($n = 62$). $^d$PCL-R = Psychopathy Checklist-Revised, coded from file as scored by clinicians ($n = 43$). $^e$HCR-20 = Historical Clinical Risk Management-20, coded from file as scored by clinicians ($n = 8$).